DOCUMENT RESUME

ED 190 591 TM 800 067

AUTHOR	Quinto, Frances; McKenna, Bernard
TITLE	Alternatives to Standardized Testing.
INSTITUTION	National Education Association, Washington, D.C. Div. of Instruction and Professional Development.
PUB DATE	77
NOTE	45p.
AVAILABLE FROM	National Education Association, Instructional and Professional Development, 1201 16th Street N.W., Washington, DC 20036 (free).
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Classroom Communication; Criterion Referenced Tests; Elementary Secondary Education; *Evaluation Methods; Higher Education; Interviews; Open Enrollment; Parent Teacher Conferences; Performance Contracts; *Standardized Tests; *Student Evaluation; Teacher Made Tests; *Testing Problems
IDENTIFIERS	National Education Association

ABSTRACT

The National Education Association (NEA) opposes the use of standardized, norm-referenced tests of intelligence, achievement, and aptitude, citing problems with their substance, structure, administration, and use. NEA suggests five alternatives: (1) performance contracts; (2) teacher-student and teacher-parent-student interviews; (3) teacher-developed tests; (4) criterion-referenced tests; and (5) an open admissions policy in higher education. Performance contracts are valuable if students understand they are working for themselves and not for the teachers. Parents should be involved and teachers should be formally trained in interviewing techniques. Teacher-developed tests, unlike standardized tests, can be related directly to course objectives and instructional materials. The purposes of such tests should be pre-determined, and formats should vary to include multiple choice, essay, and longer problem-solving tasks. Criterion-referenced tests have conceptual advantages over norm-referenced tests, and teachers must be involved in the necessarily complex process of selecting or developing objectives. Arbitrary standards, such as minimum competency levels, cutting scores, and pass-fail points can make these tests resemble normed tests. Open admissions and work-experience credits, although not widespread practices, have shown promising results. (CP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED190591

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Alternatives To Standardized Testing

TM800 067

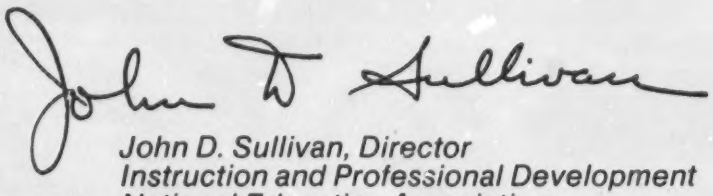
"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY
G. Felton
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

information for professional excellence

The National Education Association is committed to the goal of Professional Excellence, a commitment that can be realized only through well-informed members who ultimately will take the necessary, concerted action to achieve this goal. An important activity of the Professional Excellence program, therefore, is to provide reliable and useful information to support the practice of teaching and to enable Association and affiliate leaders to influence developing public policies in education. Accordingly, documents such as this are prepared by the NEA Instruction and Professional Development (IPD) staff, which has the major responsibility for organizing and carrying out Professional Excellence program activities.

Your comments are invited on this and other IPD documents. Also, your suggestions of other topics for future consideration will be welcome.

For more information about the Professional Excellence program, write or call Instruction and Professional Development, National Education Association, 1201 16th Street, N.W., Washington, D.C. 20036. Phone: (202) 833-4337.



John D. Sullivan, Director
Instruction and Professional Development
National Education Association

JAN 1980
RECEIVED
ERIC/TM

Alternatives To Standardized Testing

A Handbook by

Frances Quinto and Bernard McKenna
Professional Associates
NEA Instruction and Professional Development

NATIONAL EDUCATION ASSOCIATION
1201 16th Street, N. W., Washington, D. C. 20036

First Printing, August 1977
Second Printing, October 1977
Third Printing, February 1979

Copyright 1977
National Education Association of the United States

CONTENTS

Preface 5

A Point of View on Student Evaluation and Testing . . . 7

Introduction. 11

Contracts with Students 17

The Interview: Teacher-Student Conferences and Teacher-Student-Parent Conferences. 27

Teacher-Made Tests: The Paper-and-Pencil Type. 33

Objectives-Referenced (Criterion-Referenced) Tests. . . 39

Open Admissions 45

PREFACE

This monograph is a result of a final recommendation of the National Education Association's Task Force on Testing. The Task Force over a period of three years studied the issues surrounding testing, gathered data from across the country, and sought advice of knowledgeable persons and agencies within and outside the profession. As it finished its work and came to the conclusion that the "intended purposes of testing can be accomplished... through a variety of alternatives to tests," the Task Force stated that an unfinished document entitled "Some Potential Alternatives to Standardized Tests for Evaluating Student Progress and Diagnosing Learning Needs" held promise as a handbook for teachers.

The present monograph represents a next stage in developing the handbook. It contains five alternatives to standardized testing for evaluating student learning progress. There are many others, so the document may continue to be unfinished for some time to come, as suggested in the introduction. It will be the teachers of the nation who contribute the most when there is to be another edition, and another and another.

A POINT OF VIEW ON STUDENT EVALUATION AND TESTING

A resolution of the 1977 Representative Assembly delineates the National Education Association's position on testing:

> The National Education Association strongly encourages the elimination of group standardized intelligence, aptitude, and achievement tests. (Resolution 77-67)

The following statement clarifies and elaborates the Association's posture toward student evaluation generally and testing particularly.

* * * * *

The National Education Association recognizes the need of teachers, students, parents, and other decision makers for frequent and comprehensive evaluation of student progress. It is committed to excellence in the evaluation of students' interests, aspirations, and performance in all aspects of their school experience.

The NEA believes the major purposes of student evaluation should be (a) to help students choose more wisely among options and to learn better, and (b) to assist teachers in selecting the most appropriate learning experiences for individual students. The use of evaluation for categorizing and tracking students and for denying them educational opportunities should be deemphasized in favor of constructive, diagnostic, helping purposes.[1]

The NEA believes that a broad range of systems and procedures should be used for accomplishing evaluation purposes. Teachers' professional judgment should play a major role in assessing student progress. The students themselves can also be a rich resource for evaluation, with reasoned appraisal of both their own progress and that of their peers. Actual performances of students and the

---

[1]See "Report of the NEA Task Force on Testing." Reports of Committees, Councils, and Task Forces, 1974-75. Washington, D. C.: National Education Association, 1975. pp. 79-80.

products of their work are useful. So are teacher-student and teacher-student-parent conferences. There are others: simulation, contracts, interviews, narratives, and the like.[2]

The NEA believes that some tests are useful in determining student needs and evaluating student progress. Individual diagnostic instruments can be helpful if they are carefully developed, and are applied and interpreted by those who have been specifically educated for that application and interpretation. Objectives-referenced (criterion-referenced) tests, if developed to evaluate the accomplishments of specified curricula and to assist in identifying objectives yet to be accomplished by individual students, can also be useful. But as these have been constructed in some quarters, they perpetuate many of the same deficiencies characteristic of norm-referenced tests.[3] Teacher-made tests--objectives-referenced type measurements which elicit the information teachers require for planning and adapting learning activities--represent another useful alternative.

But the evidence from long experience with the most often used norm-referenced, group-administered intelligence, aptitude, and achievement tests leads the NEA to stand for their total elimination as any part of the student evaluation process, or for decision making about programs, evaluating school staffs for general policy making, or allocating resources.

In brief:

---

[2]See "Alternatives to Standardized Tests for Evaluating Student Progress." Unpublished paper. Washington, D.C.: National Education Association, 1976.

[3]See "Guidelines and Cautions for Considering Criterion-Referenced Testing." Washington, D. C.: National Education Association, 1975.

The NEA supports--

- Frequent and comprehensive evaluation of student learning progress
- Evaluation of students mainly for the purposes of helping them to make wise choices and to learn better, and of helping teachers to better understand student learning problems and to prescribe for alleviation of those problems
- The use of a broad range of systems and procedures for evaluating student learning progress, including some types of tests
- The use of some individual diagnostic instruments, carefully developed criterion-referenced tests, and teacher-made tests.[4]

The NEA opposes--

- Reliance on any single system, process, or instrument for making decisions about students' learning potential and progress
- Emphasis on the use of evaluation instruments, particularly tests, for tracking students and for denying them educational opportunities (While some kinds of special educational opportunities are appropriate for some students during part of the school day or week, students should not be grouped homogeneously by subject, class, or school.)
- The use of tests for determining the allocation of school funds
- The use of any measures of student progress to evaluate teacher competence
- The use of the National Teacher Examinations for any of the following: teacher licensure, renewal of license, selection, assignment, evaluation, retention, salary determination, promotion, transfer, tenure, dismissal.

---

[4]"Report of the NEA Task Force on Testing," op. cit.

INTRODUCTION

The classroom teacher used to be the sole authority for making judgments about students' progress in school. It was the responsibility of the teacher to assess the academic achievement of students. The community, students, and parents valued the judgment of teachers, whose professional training and experience qualified them to determine grades, promotions, retentions, and possible courses of study for the future and to evaluate overall a student's standing and progress in school.

In recent years there has been a growing move to substitute an external authority for judging student progress through the use of group standardized tests. The tests, developed mainly by commercial publishers, produce a single numerical index which is used for several purposes. It is used to compare student with student and groups of students with other groups. Sometimes it is used to compare and rate schools within districts or to compare one school district with another, a whole state with others, or one region of the country with another. It has even been used to evaluate the effectiveness of a teacher's performance without regard for the many other conditions that affect learning. The single score is also used to judge an individual student's potential ability and academic achievement. In addition, it is used by administrators and public boards to determine educational policy and to comply with federal guidelines that require testing as a basis for allocating federal funds.

This reliance on test scores for judging and comparing ability, aptitude, and achievement has also led to policies on formulating decisions for tracking or placing students in designated groups and courses of study. Using the results of tests to track students has been singled out as a way to either make possible or deny educational opportunity. There is also an increased

dependence on standardized test scores to determine grade-to-grade promotions and whether or not a student is qualified to be graduated from high school

We believe these uses of standardized test scores to be indefensible. The effects will be dealt with later in this introduction. But first we call to attention matters related to the substance and the structure of norm-referenced, group-administered standardized intelligence, aptitude, and achievement tests. These bring us to the conclusion that such tests do not have a rightful place in public elementary and secondary education, even if the problems of usage could be corrected, which is unlikely.

Problems with the Substance and Structure of Tests

The testing industry (about half a dozen large conglomerates manufacture most tests) argues that there's not much wrong with standardized tests, only their use. This argument just isn't accurate. It is true that there are large problems with the administration of tests, analysis of findings, and use of test results. But there are even larger problems with the content and structure of the tests themselves.

For example, standardized tests measure cognitive learnings almost totally. Thus their content leaves vast and important affective and psycho-motor learning areas almost completely untouched. This doesn't mean that cognitive learning is unimportant. Everybody in education -- and most people outside of it -- hope for children to be able to read and to comprehend basic mathematical concepts and do practical computations. Every effort should be made to teach all students the basic skills of reading and mathematical computations so necessary for functioning as workers and citizens in the society. But evaluating only limited cognitive learnings out of the many other important goals for which schooling is commonly held accountable results in placing undue instructional

measure, that are quantifiable. Such emphasis, it has been shown, warps the educational offering toward these limited areas, even to the point of teaching to tests. Consequently, large and important segments of education in the social and behavioral sciences, the arts, even the physical sciences are likely to be given short shrift.

Because the tests are based heavily on limited cognitive tasks (particularly recall), the ability to analyze, synthesize, draw generalizations, and make application to new phenomena are little measured by most standardized tests.

Then there is the heavy emphasis on multiple-choice items. Students who think most penetratingly and creatively can frequently justify more than one plausible answer for these. It is almost impossible to design multiple-choice items of significance that do not contain some degree of ambiguity.

Another serious problem with standardized test content is that it doesn't measure much of what schools are teaching well to all or almost all students. Such items are removed from the tests in order to provide an appropriate distribution of easy and difficult items for statistical norming purposes.

There is a laundry list of other problems with standardized tests: items irrelevant to things students know or care about, illogical items, meaningless items, complex language, and obscure vocabulary.

Finally, standardized achievement, aptitude, and intelligence tests are culturally biased. They are constructed mainly by middle- and upper-class whites and are biased in favor of white middle-class culture and values. Their content reflects this. There may be a lot right about middle-class culture and values, but in a pluralistic society the question needs to be raised: How much of such culture and how many of such values are appropriate to all segments of the society and thus should be reflected in educational measurement and evaluation?

The test makers have not turned their attention much to answering this question.

It hardly needs to be said that the tests are biased against those for whom English is a second language. Native Americans and those of Spanish and Asian descent particularly are victims in this respect, but so are all others who are admitted to the United States from countries whose language is not English. Publishers' attempts to adapt the language of the tests have not eliminated the cultural biases and values.

Problems with Test Usage

The problems of administration and of analysis, reporting, and use of standardized test results are sometimes as severe as those of substance and structure. Among them are the following:

1. The need for speed, which makes the testing situation artifical
2. Administration to students without their understanding the test purposes and uses
3. Administration under poor physical conditions
4. Use of inordinate amounts of instructional time for administration
5. Quantification of results in a way that assures one-half of the students will be below average no matter how well they do
6. Scaling in ways that result in only two or three wrong items moving a student several percentile points down the scale and thus denying him or her admission to further educational opportunities
7. Results that are not useful to teachers for diagnosing learning difficulties or planning for improving instruction
8. Results that cannot be used to help students learn better
9. Results that don't help parents understand student progress and learning problems.

We believe with Jerrold Zacharias[1] that the profession should not "retreat

[1]Zacharias, Jerrold. "The trouble with I.Q. Tests." National Elementary Principal 54: 29; March-April 1975.

to catch phrases like 'I know these tests are not very good, but they are all we have.' There are many other ways to assess a child's general competence. They may not look numerical or scientific but they are known to every teacher...."

So it is with the alternatives presented in this monograph. Most or all of them are likely known to every teacher who will read it. And most of them are not very new. What we have done is put together some of the most promising alternatives from which teachers may choose and which they may call to the attention of others who foist on them inappropriate and dysfunctional instruments for evaluating student progress.

We hope that this monograph reaches large numbers of teachers and that we will be deluged with their suggestions for adaptation and improvement of the ideas presented here as well as for additional alternatives. If this document is of some service to the profession, a successor to it based on the responses it triggers should be infinitely better.

CONTRACTS WITH STUDENTS

Contracts with students as an alternative to standardized tests can serve effectively the same purposes as the tests purport to serve. They can help to ascertain how well students are progressing in academic areas, inform the students of their achievement levels, and give them insights into their own work habits. Based on the completion of contract items, the teacher can diagnose and evaluate a student's level of competence and then designate a new program or suggest other options to help improve the student's progress. In this way, the contract can be a useful diagnostic tool for both teacher and student who have agreed to use this approach for their mutual benefit.

A Definition

Educators have adopted from the private sector the concept of writing contracts, in this case to serve an educational purpose. A contract, which is a meeting of minds between parties, consists of an offer and an acceptance. When used in schooling, the two conditions are met as student(s) and instructor agree on particular course objectives to be fulfilled by the student(s) and the support which is to be provided by the instructor. The contract provides a form of record keeping that documents student achievement in relation to specified objectives.

In a contract there must also be "consideration" -- the benefits received in return for achievement. If there is no consideration to any party, then there is no valid contract. The agreed-to consideration or gain for a student can vary, such as a specific grade, promotion to another class or group, or an indication of readiness for more difficult levels of work.

A contract should spell out desirable conditions for work and realistic and achievable goals for the student(s). This is one way to help students

believe in their own ability to perform well and to promote their image of self as an achiever.

Positive Aspects

Educators point out that writing contracts with individual students relates to the goals of schooling and produces benefits that are likely to be lasting.[2] Writing a contract with a poorly motivated student helps to improve that individual's self-image as a decision maker because the student is involved in the selection of activities. Completion of the agreement means the student has had a chance to become a partner in the learning process and, in the long term, is encouraged to undertake self-evaluation. It is a chance for the student to gain an understanding of the meaning of accountability.

In a study of disadvantaged junior high school students it was found that "grades of students under contract increased during the semester the contract was applied, whereas the grades of a similar control group declined slightly."[3]

There appear to be few restrictions or limitations for using the contract approach. Contract writing may involve a small group or an entire class. The binding contract, however, must be with the individual. If broad contracts are developed, they then must be tailored to meet individual needs and abilities.

Contracts in different areas of study can be designed to achieve one or several tasks. This, coupled with the time provided for completing a contract, will determine its content and scope.

---

[2]Thomas, G. Patience, and Ezell, Betty. "The Contract as a Counseling Technique." Personnel and Guidance Journal 51: 27-31; September 1972.

[3]Williams, Robert L., and Anandam, Kamala. "The Effect of Behavior Contracting on Grades." Journal of Educational Research 66: 230-36; January 1973.

At some levels of education, students are able to select from several alternatives the requirements they want to fulfill to meet a particular grade for a course. This has come to be known as the "grade contract." The previously cited study of junior high school students points to positive aspects of the grade contract approach, namely, it has been known to diminish the anxiety about grades for a particular course of study and thus students have had an opportunity to succeed at a self-selected level of performance.[4]

A significant conclusion was reached by investigators of a postsecondary situation, where achievement tests were in question: independent projects had more meaning for students than did tests.[5] Students who had selected and contracted for a grade level rather than taking a test viewed the test as an end product that would not contribute to their learning but only hinder independent study.

Most teachers have felt that the key to success of a contract program is the teacher-student conference. Upon completion of a contract, teacher and student can cooperatively evaluate the work, choose new assignments, and seek to clarify previous or newly stated objectives.

Some Cautions

There are some cautions to consider in using contracts. First, students should understand that they are producing a product for themselves, not performing an activity for the instructor. A contract should be utilized as a process for learning rather than merely as an instrument for getting a job done. "The process, not the instrument, is the key to developing successful contracts

---

[4]Ibid.

[5]Poppen, William A., and Thompson, Charles L. "The Effect of Grade Contracts on Student Performance." Journal of Educational Research 64: 420-24; May-June 1971.

with students...[it] offers a unique opportunity to personalize and humanize instruction."[6]

The process of contracting with students should allow for renegotiation if contract terms become a hazard or are viewed as unreasonable by either the teacher or the student. One hazard may occur when a contract which was acceptable at first appears too complex for the student to deal with when working alone. Unanticipated difficulties can increase as time passes, resulting in either loss of student interest or undue frustration and, as a result, the student may give up. Another hazard occurs for some students who have a poor self-image or who are poorly motivated. Given a choice, they may select objectives of minimum difficulty even though they are capable of performing more complex tasks.

Grade contracting also should allow either students or teachers to renegotiate. When during the prescribed period a student feels he/she is producing a product above the agreed-to grade level, in some instances an alternative product should be acceptable. As an example, if a written description of a student's observations on a field trip was the contract requirement, the student or the teacher might want to renegotiate and substitute a taped report or another product which could demonstrate the student's ability as an observer.

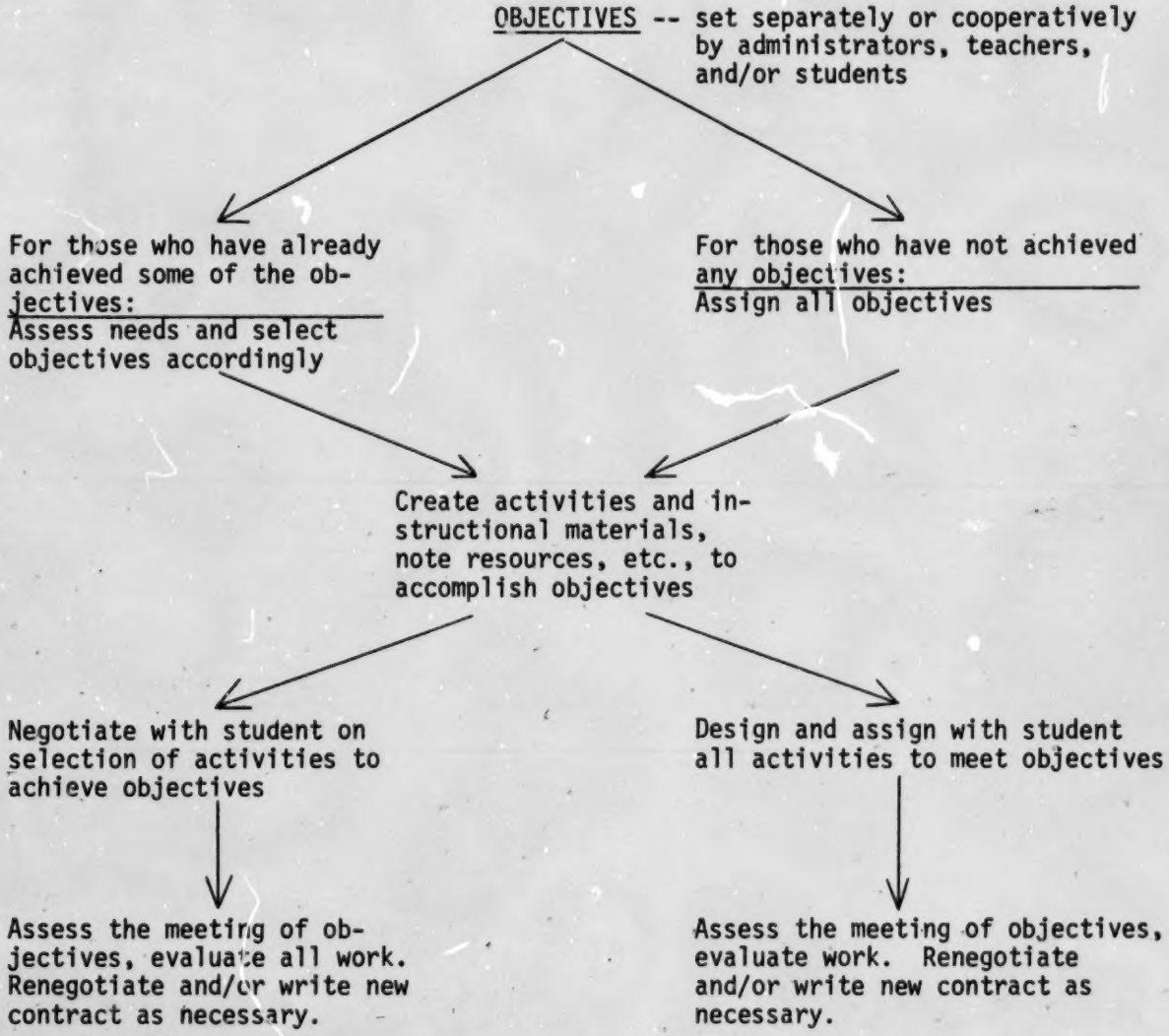
It should be understood that the use of contracts is not a laborsaving device. In all probability it requires more time and effort than do most conventional programs. The most successful contract programs take place when objectives are clearly thought through and understood by all parties before tasks are offered and accepted. The contract plans that stress too much the

---

[6]Bockman, John F., and Bockman, Valerie M. "Contracting for Learning Outcomes: Potential and Limitations." Bulletin of the National Association of Secondary School Principals 57: 17-26; February 1973.

signing and writing of legalistic-type documents (often standardized and imposed) have little to offer toward personalizing and humanizing education.

One teacher has noted that the contract approach is most successful in an open classroom setting.[7] He suggests using the following procedure, allowing for flexibility in time in order to achieve the objectives.


OBJECTIVES -- set separately or cooperatively by administrators, teachers, and/or students
For those who have already achieved some of the objectives:
Assess needs and select objectives accordingly
For those who have not achieved any objectives:
Assign all objectives
Create activities and instructional materials, note resources, etc., to accomplish objectives
Negotiate with student on selection of activities to achieve objectives
Design and assign with student all activities to meet objectives
Assess the meeting of objectives, evaluate all work. Renegotiate and/or write new contract as necessary.
Assess the meeting of objectives, evaluate work. Renegotiate and/or write new contract as necessary.

[7]From a conversation with Joseph Czarnecki, member of the Maryland State Teachers' Association.

Instructors who practice traditional procedures but who want to apply the contract method should not abandon all structures that have been successful for them over a period of time. The contract should serve as a communications instrument about objectives and tasks, not the determiner of class organization. As a beginning, time blocks for contract work should be set by the teacher and choice of subject may be basically the student's decision. More freedom of choice can be granted as the first steps are mastered by both student and instructor.

Contract Samples

On the following pages are samples of three types of contracts for sixth-grade students. The number of objectives and activities will vary depending on student needs. In some instances, the contract items may be done in any order and selected to meet student needs or objectives.

(The Nervous System)

OBJECTIVES - Can you...

1. Name the 7 stimuli (things that our nerves can feel) that affect our nervous systems?
2. Explain what nerve fibers are?
3. Label parts of a neuron?
4. Explain a nerve impulse, stimulus, synapse?
5. Describe the Central Nervous System and the Autonomic Nervous System and the differences between them?

ACTIVITIES (designed by teacher to answer objectives)

Drawing #1 - Neurons, page 251
Drawing #2 - Sensory Receptors, page 253
Drawing #3 - Motor End Plates, page 255
Drawing #4 - Reflex Arc, page 256

Answering Objectives

1. Answer the objectives on separate sheet.
2. Answer each objective in a complete sentence.

Learning Centers (materials can be developed by teacher and other resources noted)

Do each of these Learning Centers only after doing the drawings.

1. Neurons
2. Sensory Receptors
3. Motor End Plates
4. Reflex Arc

Signature_____

(Study of the Earth)

NAME:_____ Date Started:_____
Date Completed:_____

OBJECTIVES [limited sample] - Can you...

1. Describe the atmosphere and lithosphere?
2. Name and explain the layers in the lithosphere?
3. Define a mineral (218)?
4. Explain the origin of igneous, metamorphic, and sedimentary rocks?
5. Name the 5 elements that make up 91% of the earth's crust

DIRECTIONS

1. Each book assignment question must be answered as a complete statement. (Example: change question 1 from, Explain what an atmosphere is, to The atmosphere is....)
2. Skip a line between statements.

Learning Centers

Use materials and books and do the following as they relate to objectives.

1. Layers of the Lithosphere
2. The Atmosphere, Lithosphere, and Hyrdrosphere
3. The Composition of the Earth's Crust

Group Activities

1. Movie,"Rocks That Form on the Earth's Surface"
2. Class on "Identifying Rocks by Physical Tests"

GENERAL GRADE CONTRACT

(Postsecondary)

I, _____

agree to complete _____
(No. of tasks or other details)

as per the attached assignment sheet by _____
(Date)

for the grade of _____.

I understand that upon completion of the assignment a conference will be held to determine the attained grade.

Failure to produce the agreed work at the proper time and at the specified quality results in a breach of contract.

_____ _____
(Student signature) (Instructor signature)

_____
(Date)

THE INTERVIEW: TEACHER-STUDENT CONFERENCES AND TEACHER-STUDENT-PARENT CONFERENCES

Ginsburg[8] reports, "To know a child, we must talk with him in depth, going far beyond what can be learned from a series of standard questions." The maker of this statement wasn't applying it to affective areas such as attitude development or citizenship learnings. He was talking about arithmetic. And the "standard questions" referred to could well be standardized tests. Direct dialogue with students in the form of interviews can accomplish many things that can be claimed for no standardized test.

Evaluating student progress through direct oral communication with students has been around as long as the other alternatives discussed in this monograph, longer than most. And it has been used successfully and recommended by some of the leading thinkers about education, including Jean Piaget. Some teachers have made good use of this technique over a long period. But interviewing students to ascertain their learning level, to identify learning problems, and to assist in determining appropriate remediation isn't easy. To become so skilled requires considerable understanding of the technique and much practice. Teachers need opportunities for formal learning in such skills in either preservice or in-service education.

The interview as an evaluation technique can be as broad or as limited as the evaluator wishes to define it. As suggested by Ginsburg, it can be as specific as a teacher sitting with a child and asking him or her to share information about how to solve particular mathematics problems. Or it can be as broad as the technique developed at the North Dakota Center for Teaching and

---

[8]Ginsburg, Herbert. "Talking with Children: An Alternative to Testing." Educational Technology 15: 41; April 1975.

Learning which probes one or more of the following: classroom activities; materials; scheduling arrangements; organization; diagnostic and student evaluation approaches; student-peer interaction; student-teacher interaction and relationships; classroom problems, difficulties and high points; the teacher's goals; the teacher's reaction to some general school goals as related to his or her own experience in the classroom; the teacher's relationship, experiences, and interaction patterns with parents; perceptions about the community environment in which the classroom operates; and changes in the classroom in the course of the year in all of the preceding.

As with other forms of evaluation, whether or not to use the interview, how extensively, and in what form depends on both the educational and the evaluational purpose at hand.

One purpose which can be served by a well-managed interview is to determine how a student approaches learning and problem solving. Through such a procedure teachers can discover much about students' individual learning styles and thus make better decisions on what instructional methods to use with particular youngsters.

Another purpose is to probe the appropriateness of the classroom climate. Teachers can learn from individual one-to-one conferences what things in the social, organizational, and physical setting of the classroom may be hindering the learning of particular students.

If the teacher uses teacher-made tests (dealt with in the next section), individual interviews based on test items to determine the kind and order of reasoning that caused particular students to respond as they did can serve the diagnostic purpose of identifying faulty learning processes. Such a procedure can also help identify illogic and ambiguity in the tests. James has made a good case for one-to-one evaluations of this nature:

> If we need to test (rather than simply observe daily behavior) we should be looking for forms of testing which enable the child to tell us, directly or indirectly, "This is the way I perceive the problem, this is how I perceive it in relation to me, and this is how I perceive myself able (or unable) to act in relation to it." So long as we test in ways which divide cognition from affect and both from will we shall not have evaluation procedures that are worthy of open education.[9]

As James suggests, the interview can be a highly individualized and personalized tool for evaluating student learning progress. One way of personalizing it even more is to involve parents.

Parent Involvement

Almost all parents are interested in their children's progress in school. While there are problems in getting parents to come to school to meet with their children and their children's teachers, once these have been surmounted, the rewards of such activities are great. In a constructive, supportive situation in which parents, student, and teacher jointly discuss and appraise the student's learning progress, new insights can be gained on successes, potential, and needed remediation.

In the first place, parents bring to the interview situation unique ways of looking at their child's progress. In a counseling setting they are frequently able to verbalize about or help the student describe his or her own learning styles and problems.

Such involvement can serve an additional important purpose. It can cause parents to take renewed interest in their children's learning, to be better able to encourage and help them at home, to reinforce behaviors that appear most promising and discourage those that are deterrents to learning. If the contract

---

[9]James, Charity. "Toward a Shared Appraisal." Evaluation Reconsidered: A Position Paper and Documents on Evaluating Change and Changing Education. New York: Workshop Center for Open Education, 1973. p. 13.

plan is in use as an evaluation technique (see the previous section), the parent-teacher-student interview is a good opportunity to get the parents to understand the implications of the contract and the student's progress and to help the student develop appropriate commitment to fulfilling the contract.

Parent-teacher-student interviews should make use of the broadest range possible of student work samples for developing parental understanding of student status and progress. The tripartite nature of the parent-teacher-student interview makes it even more necessary for teachers to have formal preparation in the process.

Preparation for Conducting Interviews

Teachers who are planning to use or are already using the interview technique without formal introduction to it will do well to get specific instruction, including simulated and real experience in interviewing. There is much to be learned about establishing a climate of mutual trust and security; patterning the total process to elicit all the data that is wanted; asking questions in ways that will result in full disclosure of attitudes, values, and uncertainties about all things being evaluated; how to talk less and gain longer narrative responses from interviews.

Opportunities for teachers to learn the interview technique need to become broadly available if it is to be used constructively and positively. There currently are few or no such opportunities in either preservice or in-service education, but a variety of sources for such training are available. Nearest to the instructional setting itself are school counselors and school psychologists, many of whom have had sound preparation and extensive experience in using the interview for appraisal purposes. With some joint planning among these specialists and classroom teachers, in-service sessions can be carried out which will prepare teachers to conduct evaluation interviews. Beyond the school

there are agencies specializing in organizational development (such as the National Training Laboratories) which are prepared to conduct interview training. In addition, the personnel departments of many businesses and industries have staff members skilled in interviewing for personnel selection. With some modification in their techniques and with advice from teachers, they could soon be able to conduct successful interview training sessions for teachers.

TEACHER-MADE TESTS: THE PAPER-AND-PENCIL TYPE

Some teachers believe that students become test weary; that is, as the number of tests increases, efforts to achieve decrease. The reasons range as follows: Students either do not know or do not understand their scores and so feel they have no real stake in doing well. Where students are given the results, the time lapse between taking the test and receiving the results is too long to sustain interest; by that time the students have progressed to more complex levels of study. Since the general practice is to not review the exercises, taking the test becomes another impersonal activity performed for no one in particular and care is not used in answering items.

One remedy for this type of situation is the teacher-made test. Tests devised by teachers can help determine what machine-scored, group-administered tests cannot. They can be tailored to specific situations and individual needs. Teachers are professionally prepared and competent to correlate a number of essential factors about their students which should be taken into account in evaluating achievement: the milieu in which a student exists, both in and out of school; individual variations and unique qualities; changes in attitudes and motivation as observed over a period of time. They can relate tests directly to the course work, to their objectives, and to the instructional materials used with students. They can review written test exercises soon after administering the test, noting how the students responded and why they responded in a particular way.

Developing Teacher-Made Tests

Developing written tests for classroom use is a routine activity for most teachers. These tests, a form of the objectives-referenced type (discussed in the next section), help to serve the needs of both students and teachers. They disclose where a student stands in relation to classroom objectives and are a

guide for providing appropriate resources and help. As noted by Richard L. Venezky, "To sanction periodic testing solely for the purpose of curiosity...is to encourage inefficient and irresponsible management."[10] Because of the potential information that teacher-made tests can offer, they should be carefully developed rather than being "off the top of my head" activities.

What are some elements that can help to make paper-and-pencil teacher-developed tests effective and reveal the kinds of information wanted?

Before developing a test, decide what purpose it will serve and how the results will be used. A test may be administered to determine the strengths or weaknesses of a group or individual or class knowledge of the subject. It can be given either at the first meeting of a class, prior to the introduction of a unit of study, after completion of a unit or course, or as a reinforcement activity; or it can be the "every Friday" test which helps monitor progress in orderly fashion. Another purpose is to assist in the identification of students who need specialized help and the areas in which help is needed. It can also be a means to observe and identify special talents. Inferential thinking should be tested only after a student has mastered basic information or sources from which inferences will be made.[11]

A first step in test development is to list the kinds of information to be elicited and then decide the best format for securing the information. The content of the test is determined by the subject area, the kind of information and activities that have been provided for the students, and the amount of time allotted for the test. A decision should be made as to the number or percent

---

[10]Venezky, Richard L. Testing in Reading. Urbana, Ill.: ERIC Clearinghouse on Reading and Communication Skills and National Council of Teachers of English, 1974. p. 6.

[11]Ibid., p. 9.

of items for each category. The format should include problem solving and application of information to new situations. Developing a grid such as the one illustrated below can be helpful.

Format for Test

Areas of Information in Language Arts	Multiple Choice	Essay	Problems: Applying Information
Use of question mark	(10 min.) 4 questions	(20 min.)	(approx. 15 min.)
Developing a paragraph		1 question	
Sequential thinking			2 questions

It should be noted that multiple-choice exercises are limited to the kinds and quality of information they can provide. They measure only cognitive knowledge. Since deductive or analytic thinking is not necessary, students -- particularly those in the upper grades of secondary school -- admit to much guessing at the "correct" answer or to arbitrarily selecting an answer in the hope that it will be right. Ease of machine scoring is the justification for the multiple-choice format, but there is a burden for those who select this route: useful multiple-choice items are difficult to develop and take time to perfect. Test exercises should never be constructed so as to stump, catch, or confuse but should be stated as clearly as possible to avoid misunderstanding. The developer should be careful to assure that the correct answer is not similar in wording or meaning to an incorrect one. Nor should the right answer in one item be the wrong answer in another. This is confusing, serves no useful purpose, and takes time away from other constructive, thoughtful responses.

Essay-type test items, which take less time to develop, take much more time to review and evaluate. It would be wise to be consistent in evaluating essay responses. Weight can be given to such factors as organization of topic, description, sentence structure, and snythesizing. There should also be consistency in how much grammatical construction or punctuation influences evaluation of tests other than in the language arts. Students can have a part in deciding or selecting the criteria to evaluate essays in different disciplines.

Just as variety is a necessary ingredient in classroom activity, so too there should be variety in teacher-made tests. In addition to the general format, there can be matching of words or ideas, some completion-type exercises, true-false items, graphs, pictures, diagrams, short and long presentation of problems, items that can be worked through quickly, and those that take somewhat longer.

If reading comprehension or other reading abilities are not being tested, then questions and problems, including directions, should be presented in a form that all students will understand, or directions can be given orally.

Teachers generally accept the challenge of creating new test items, but effective items and problems should be saved from year to year since valuable tests are not simple to construct. If a test is from the "stockpile," review the items to be sure that all areas have been covered in class. And whether newly constructed or not, work through the test first, or have another instructor do it, to help refine and clarify both presentation and content.

As noted earlier in this section, judgments about student performance should be based on a variety of measures over a period of time rather than on a single source. While the emphasis here has been on the paper-and-pencil type test, appraisal should be both written and oral and should cover, for example, estimating distances, quantities, and quality; generalizing from the specific;

drawing conclusions; and synthesizing and analyzing information. Certainly a student's creative abilities in problem solving or artistic endeavors cannot be overlooked. Standardized tests, which consume so much time and money, can in no way give the teacher or other decision maker insight into the above and other important affective aspects of student achievement and potential.

OBJECTIVES-REFERENCED (CRITERION-REFERENCED) TESTS

In order for teachers and others to have a common understanding of the meaning of the alternative discussed in this section, a somewhat arbitrary definition is provided below. This definition is offered recognizing that psychometricians and others have defined separately objectives-referenced tests (ORT), criterion-referenced tests (CRT) and domain-referenced tests (DRT).[12] Our study of these definitions leads us to believe that they are overlapping and unclear and the tests likely indistinguishable operationally.

Objectives-referenced test, as used here, is defined as a test wherein all items are designed to evaluate the accomplishment of specifically stated instructional objectives. Student accomplishment is reported and interpreted in terms of how nearly the student achieved all the instructional objectives, not how it compares with the performance of others.

Two prominent advantages frequently put forth for objectives-referenced tests over norm-referenced tests are:

1. They more nearly reflect specific instructional purposes (objectives)
2. They evaluate how nearly students are achieving mastery of objectives rather than where they stand in relation to others in terms of some average or norm. (In this respect, objectives-referenced tests might be expected to serve more diagnostic purposes than the typical norm-referenced test.)

We believe these "conceptual advantages" alone make objectives-referenced tests worth pursuing.

There are several important considerations that must be taken seriously if ORT's are to fulfill their promise as viable alternatives to group-administered, norm-referenced standardized tests.

Let us examine some positive considerations and then some caveats.

---

[12]Sanders, James R., and Murray, Stephen L. "Alternatives for Achievement Testing." Educational Technology 16: 17-23; March 1976.

Positive Considerations

The Objectives. As with any evaluation, the objectives to be evaluated by ORT's need to be carefully defined and developed, to be clear, to be broadly agreed on, and to remain flexible enough so that a variety of creative approaches to their accomplishment can be selected by teachers. ORT's will be only as good and as well accepted as the objectives on which they are based.

Derivation, Locale, and Involvement in Objectives Development. The objectives themselves must be directly derived from and functions of agreed-on goals for schooling. Because it is teachers who must carry out the activities for accomplishing the objectives and who for the most part must evaluate the degree to which objectives have been accomplished, teachers must be fully involved in selecting and/or developing the objectives.

The development of ORT's should be mainly local (by the school district or by individual schools within the district). In this way they can more nearly reflect specific local objectives and methodologies.

Analysis of Agreed-on Instructional Areas. The ORT holds promise for making its highest contribution operationally if objectives are chosen by careful, in-depth analysis of the instructional areas to be tested. That is, the content of and/or behaviors in an area of instruction should be determined and learning-outcome objectives and tests should be directly related to the content and behaviors.[13] If this concept works well it should be possible to identify clearly which objectives are not being accomplished and by which students. Also, teachers would have a better grasp of the diagnostic function of teaching than they typically have had in the past. But even then, identifying which objectives particular students are not accomplishing is only half the job.

---

[13]This appears to be the uniqueness claimed by Hively and others in defining domain-referenced tests as separate from ORT's. We fail to see enough difference to justify separate definitions.

Once the unaccomplished objectives are isolated, one needs to know how to get particular students to accomplish particular objectives--the "What works with Johnny and for what purpose?" question. Even with the high level of professional expertise of teachers and the considerable research on teaching (especially in the basic skills), this remains a complex and muddled area in terms of prescribing remediation for some student learning problems. The ideal of more clearly describing objectives for student learning and better determining which objectives individual students have not learned is reasonable. But one should not expect too much in the way of a solution to every learning problem, with the state of research and development on remediation as it is.

Development of the Tests. When opportunities are provided for teachers themselves to develop the ORT's there is considerable promise that the tests will directly reflect instructional intent. At least they are likely to be more dependable in assessing individual learning outcomes and attendant instructional techniques than tests developed solely by measurement experts. But they may be technically less precise. (The caveats section deals further with this problem.)

Reporting Results. Under careful control there is potential for more constructive use being made of ORT results than is commonly the case with norm-referenced tests. Through them can be avoided single comparative scores which are so often used to categorize or pigeonhole students and deny them access to particular educational opportunities. If reporting systems for ORT's emphasize individual student accomplishment (or nonaccomplishment) of objectives, one of their important constructive purposes will be best served. (The reporting procedure for ORT's also has caveats.)

Some Caveats

We reiterate what has been emphasized before: We support the use of some testing procedures for evaluating outcomes of student learning if the procedures

are carefully developed substantively and structurally and administered fairly and reasonably, and if the results are analyzed and used with great caution and used in combination with a broad range of other criteria for assessing student learning outcomes. We believe that ORT's are among the promising testing procedures. But in some places they have been developed so hastily and incompletely and have been promoted so out of proportion to their ability to deliver on the promises that we feel compelled to provide cautions to those teachers and others who are considering ORT development and use.

Select objectives carefully. All agreed-on educational goals need to be taken into account in objectives development and selection, not just those easiest to measure or thought to be important by only a limited sector of the professional staff or by test makers who may view the objectives to be accomplished in an educational area or areas from too limited a perspective. It has been noted that test makers who wouldn't think of including items in tests without field tryout sometimes accept and discard objectives with little or no rationale.

Development is complex. Objectives-referenced tests won't be any easier to develop than norm-referenced tests. In fact, they are likely to be more difficult to develop. This does not mean teachers shouldn't take on the task. Actually they've been doing it for years. The "Friday quiz," long used by teachers to assess how well students have learned the instructional objectives of Monday through Thursday, is a near prototype of what is currently being called objectives-referenced tests. But if present-day ORT's are to have higher degrees of reliability and validity than in the past, and if they are to be more useful for diagnosis of individual learning problems, then large resources of talent, time, and money will need to be provided for their sound development. And while teachers must play a major role in ORT development, testing, and implementation,

they will want to call in other professional specialists, particularly evaluation specialists, to assist them.

Field testing is crucial. Unless ORT's are thoroughly field tested in a variety of educational settings, under varying conditions, with different populations of students, and with key variables controlled which may affect outcomes, they will be less than helpful.

Arbitrary standards should be avoided. One advantage put forth for ORT's over normed tests is that they eliminate the practice of placing all students in a range related to an average and judging all those below the average or below some point in the range to be inadequate. But ORT's don't assure the avoidance of this objectionable practice. In fact, some test developers currently are establishing minimal competency levels, cutting points, and pass-fail points for ORT's which can then be applied to individual students or used to compare individual and group achievement with the achievement of other individuals and groups in ways that are as invidious as are norms. As ORT's are developed, caution must be taken that highly arbitrary standards are not built into them.

Beware of norm-referenced tests in disguise. When arbitrary standards such as described above are established, the ORT's come to have the same effect as norm-referenced tests. Since the item banks from which some commercial test makers draw in developing ORT's are the very same from which they draw for normed tests, the establishment of arbitrary standards makes them almost identical to normed tests. Such procedures applied to ORT's obviate the major advantages claimed for them over normed tests:

1. That they judge students on individual performance rather than on how they compare with others
2. That they better reflect particular instructional objectives.

ORT's do not automatically correct cultural and socioeconomic bias. ORT's are as subject as are norm-referenced tests to bias against those of cultures different from the culture of most test makers, those whose major language is different from that of the tests, and those whose opportunities have been highly limited because of socioeconomic status. Most ORT developers have done little to eliminate bias from their tests.

Teachers should not be evaluated on ORT's -- any more than on norm-referenced tests. Any test measures only a small part of what students learn and of teacher effectiveness. Reputable researchers reject the evaluation of teaching on the basis of any outcome measure of student learning.

In Summary

If you are going to develop or use objectives-referenced tests--

1. Request that you, as the teacher closest to the learner, have major authority for establishing instructional objectives, with the advice of other professional specialists, the parents, and students, as appropriate.

2. Be sure the objectives are consistent with agreed-on local and/or state goals for schooling.

3. Get technical advice on test-item development from evaluation specialists, as needed.

4. Employ a broad range of test formats other than multiple-choice and true-false. Don't be intimidated by psychometric rhetoric that says evaluation devices must demonstrate statistical reliability and validity in order to be useful. Clearly it is helpful if some do, but the evaluation of a number of the most important educational outcomes requires broader measures which may not be highly quantifiable. Some of them are described elsewhere in this monograph.

5. Allow a considerable period of tryout of your new tests to get reactions from teaching colleagues, specialists, students, and parents, as appropriate, on how helpful they are for diagnosing learning problems and prescribing for them.

6. Keep the door open for using a broad range of devices for evaluating student learning progress. Don't get hung up on objectives-referenced tests or any other single criterion.

OPEN ADMISSIONS

The practice of open admissions is included here to illustrate that a desirable alternative to the deleterious effects of standardized tests and use of their results is to give no tests at all.

In 1973 a student study group reported on institutional responsibilities in higher education, stating that there should be for all "the right to a free higher education." The students called for the elimination of all barriers among institutions and cited as barriers admissions criteria, costs, and educational tracking, all of which "negate the [undeveloped] potential of many people."[14]

Since the time that statement was made, there has been a continuous change in public attitude toward acceptance and support of these goals set by students. There have been institutional changes, eliminating the entrance barriers to some extent and making postsecondary education available to more people.

The barrier of "admissions criteria" has been for students the scores on group-administered standardized achievement tests. The scores have been known to give the advantage to students who least need an advantage--those from families at the middle and upper socioeconomic levels--because they most often score better on achievement tests than students from the lower socioeconomic level. In high school, the more advantaged students have generally been tracked into programs and classes comprised of like students who eventually take standardized college entrance tests, score well, and then go on to higher education.

The universities that rely on standardized test scores select their students from those who score the highest, and they invariably are from

---

[14]Rosen, David; Brunner, Seth; and Fowler, Steve. Open Admissions: The Promise and the Lie of Open Access to American Higher Education. Lincoln: Nebraska Curriculum Development Center, University of Nebraska, 1973. p. 1.

upper-income groups. These universities have operated on the assumption that since standardized tests purport to predict success in college, only high-scoring students have the capability of performing well at the college level.

A View of Change

Educational institutions would have to agree that standardized tests are used mainly to exclude students from rather than include students in schools and programs. In view of this, some postsecondary schools that were overcrowded in the early '70's set up the machinery to select students for admission on criteria other than standardized tests. Selection procedures became more complex and expensive and added on admissions officers a burden they were not prepared for. Some schools began to place more value on the personal interview, on high-school class standing using the professional judgment of instructors, on grades, on student products, on artistic achievement, and on recommendations from teachers. In some instances these factors were weighted, and standardized tests began to play a diminished role. Some state colleges and universities were given the mandate to open the doors to all who wanted to attend and they set up an "open admissions" policy. Financial support was available to students who needed it and schools provided special remedial programs and other academic assistance where necessary. The effort to help students achieve academically in open admissions schools meant a redirection in responsibility for faculty, students, and the state coffer.

Open admissions can be viewed as humane and fair and as promoting the concept of equal educational opportunity. Even though it is far from being a generally accepted practice, there have been some promising results. In the fall of 1975, Michael Harrington reported in the New York Times that "CUNY [City University of New York] is making the transition toward the goal of quality, mass higher education. In fact...the rate at which students who enter the

university remain enrolled...compares favorably with a national sample."[15] The difficulties inherent in the commitment to open admissions have been, at times, insurmountable for both instructors and students, and the system has not always worked to the benefit of minority students. But more than other higher education policies, "open admissions has committed itself to the proposition that young people who are poor, of minority background or working-class can... benefit from going to college."[16]

A change in CUNY admissions policy occurred in 1976, however, because of the financial strain experienced by New York City and the growing nationwide demand for proof of proficiency from high school graduates. Students now must pay a tuition fee and must be in the upper one-third of their secondary school class. Remedial programs are provided for those whose averages are below 70 percent.

This exclusion of students from educational opportunities occurs at other levels too. A recent study of the U. S. Department of Health, Education, and Welfare has found that the testing, tracking, and other ability groupings based on test scores have created a kind of de jure segregation in Southern schools. The study points out that testing and tracking result in virtually all-black classrooms made up of children whose backgrounds started them off on the wrong foot. This gross misuse of test results works only to defeat the children's self-images and denies them access to proper learning situations, including association with higher-achieving classmates.[17] For most children who are tracked into "slow" groups, achievement on standardized tests is more a reflec-

---

[15]Harrington, Michael. "Keep Open Admissions Open." New York Times Magazine, November 2, 1975. pp. 16-19.

[16]Ibid.

[17]Mills, Roger, and Bryan, Miriam. Ability Grouping and Segregation. Atlanta: Southern Regional Council, 1976.

tion of their standing in the social structure than it is a reflection of their native intelligence.

Proceeding from Here

Education in general, and higher education in particular, has felt the sting of recent fiscal crises in the country. There has been a great reversal in the amounts of financial help available to students and to institutions. The unavailability of federal and state monies has caused a cutback at some institutions in the number of faculty, in programs, and in the number of students selected, and a tightening of the selection process.

The worthy goals of open admissions coupled with the diminished reliance on the infallibility of standardized test scores is reflected in the educational policies of some schools. Pace University in New York, for example, has a well-developed program of adult education that allows people to enter or reenter degree programs based on performance rather than on standardized achievement tests. It has opened up opportunities for people who want to pursue an education while working, and for those who have worked for several years but want now to have the benefits of college.

A recent Sunday edition of The Washington Post carried the following advertisement: "New course in The Washington Post offers college credit."[18] The course originates from the University Extension, University of California, San Diego, and is funded by the National Endowment for the Humanities. Anyone may enroll and undergraduate credit will be awarded to those who successfully complete the course requirements. The program stands as an example of the public's acceptance of an open admissions policy and alternative approaches to learning.

---

[18]"Potomac" magazine, Washington Post, August 1, 1976.

Another advertisement reads, "Plan your own college program." It goes on to explain how to get credit for work and life experience and earn from an accredited school either a B.A., B.S., M.A., or M.F.A. degree.

The Carnegie Corporation of New York reported in fall 1975 that more than 100,000 people would try that year to gain college credit based on their experiences in life[19]--proving knowledge by experience is one form of the competency-based approach to education. It is another example of the public's desire to open the doors of higher education and to discard educational testing and tracking. It also displays a move to broaden the opportunities for learning and for earning credit. It is a search to provide alternatives, to receive recognition for academic achievement, and to offer recognition for ability and accomplishment.

The approaches to open admissions discussed here are evidence that dependence on a standardized test score for admission to schools and programs denies equality of access to educational opportunity for all. Criteria for acceptance should be based, to a large degree, on student desire and motivation to enter a school and to accept its programs of study.

---

[19]"There's More Than One Way To Earn a College Degree." Carnegie Quarterly 23: 1; Fall 1975.